

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-143193

(43)Date of publication of application : 29.05.1998

(51)Int.Cl.

G10L 3/02

G10L 3/00

G11B 20/02

(21)Application number : 08-296104

(71)Applicant : MATSUSHITA ELECTRIC IND CO LTD

(22)Date of filing : 08.11.1996

(72)Inventor : YAMASHITA SACHIHIRO

GOTO SHOICHI

TANIGUCHI SHUHEI

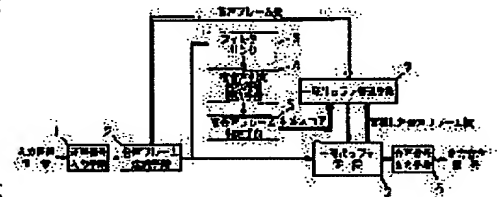
ISHIZU ATSUSHI

## (54) SPEECH SIGNAL PROCESSOR

## (57)Abstract:

**PROBLEM TO BE SOLVED:** To suppress the absence of a speech signal containing an important key word and enable easy-to-hear fast-forward speech reproduction by providing a temporary buffer managing means which discards speech frames flexibly.

**SOLUTION:** A filter bank 3 divides inputted speech frames into frequency bands by using a filter bank equipped with plural band-dividing filters. A voiced sound decision object band selecting means 4 which inputs the band-divided speech signal selects the speech signal and inputs it to a voiced sound frame decision means 5. The voiced sound frame decision means 5 compares the spectrum amplitude level of a band-limited speech signal with a threshold level. A temporary buffer managing means 7 rearranges management information in the decreasing order of decided score values by referring to the decided score values for the respective speech frames in the temporary buffer management information. After the rearrangement, specific speech frames are selected in the decreasing order and left in a temporary buffer means 6, and the remainders are discarded.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-143193

(43) 公開日 平成10年(1998) 5月29日

(51) Int.Cl.<sup>6</sup>

識別記号

F I

G 1 0 L 3/02

G 1 0 L 3/02

A

3/00

5 1 5

3/00

5 1 5 D

G 1 1 B 20/02

G 1 1 B 20/02

G

審査請求 未請求 請求項の数 2 O L (全 10 頁)

(21) 出願番号 特願平8-296104

(22) 出願日 平成 8 年(1996) 11 月 8 日

(71) 出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72) 発明者 山下 祥宏

大阪府門真市大字門真1006番地 松下電器

産業株式会社内

(72) 発明者 後藤 昌一

大阪府門真市大字門真1006番地 松下電器

産業株式会社内

(72) 発明者 谷口 周平

大阪府門真市大字門真1006番地 松下電器

産業株式会社内

(74) 代理人 弁理士 滝本 智之 (外 1 名)

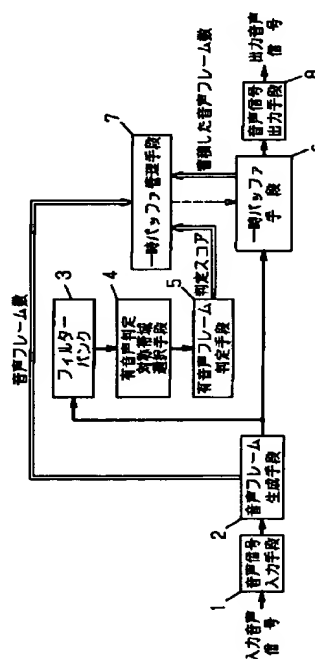
最終頁に続く

(54) 【発明の名称】 音声信号処理装置

(57) 【要約】

【課題】 重要キーワードの欠落の少ない、聴き取りやすい早送り音声再生を可能とする音声処理装置の提供を目的とする。

【解決手段】 音声信号入力手段と、音声信号を分割する音声フレーム生成手段と、音声信号を任意の周波数帯域幅に分割するフィルターバンクと、周波数帯域分割した音声信号から、任意の周波数帯域を含む音声信号を選択する有音声判定対象帯域選択手段と、選択した周波数帯域制限の音声信号から、有音であるか無音であるかの判定を行なう有音声判定手段と、音声信号を一時蓄積する一時バッファ手段と、有音であるか無音であるかの判定結果と一時バッファ手段に蓄積された分割音声信号の数をもとに、一時バッファ手段に蓄積した音声信号を分割音声信号毎に管理する一時バッファ管理手段と、音声信号出力手段を備えたことを特徴とした音声信号処理装置。



## 【特許請求の範囲】

【請求項 1】 音声信号を入力とする音声信号入力手段と、前記音声信号入力手段からの音声信号を分割する音声フレーム生成手段と、前記音声フレーム生成手段によって分割した音声信号を任意の周波数帯域幅に分割するフィルタバンクと、前記フィルタバンクによって周波数帯域分割した音声信号から、任意の周波数帯域を含む音声信号を選択する有音声判定対象帯域選択手段と、前記有音声判定対象帯域選択手段によって選択した周波数帯域制限の音声信号に対し、有音であるか無音であるかの判定を行なう有音声フレーム判定手段と、前記音声フレーム生成手段からの音声信号を一時蓄積する一時バッファ手段と、前記一時バッファ手段に蓄積された音声信号を、前記有音声フレーム判定手段での判定結果と前記一時バッファ手段に蓄積された分割音声信号の数をもとに、分割音声信号単位に管理する一時バッファ管理手段と、前記一時バッファ手段からの音声信号を出力する音声信号出力手段を備えたことを特徴とした音声信号処理装置。

【請求項 2】 前記有音声フレーム判定手段での判定結果から、前記音声フレーム生成手段からの有音声信号を選択する有音声フレーム選択手段と、前記一時バッファ管理手段の代わりに、前記一時バッファ手段に蓄積された分割音声信号を、前記一時バッファ手段に蓄積された分割音声信号の数をもとに、分割音声信号単位に管理する一時バッファ管理手段を備えたことを特徴とした請求項記載 1 記載の音声信号処理装置。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】 本発明は音声信号の早送り再生に関するものである。

## 【0002】

【従来の技術】 最近のビデオ機器などでは、映像の早送り再生技術の発展とともに、音声の早送り再生技術も発展してきている。

【0003】 映像の早送り再生は、視覚的に気にならない程度のコマ落としを行い、映像を間引きながら再生することで早送り再生を可能としている。

【0004】 一方音声は、時間データの間引きを行い、時間長を短くすることで早送り再生を可能としている。この時間データの間引きは、単純に音声データの内容に関係なくある一定長を間引く方法と、無音音声部を検知し、これを間引く方法がある。またこれらの 2 つを合わせた方法もある。

【0005】 従来例として、単純間引きと、無音音声部を検知しこれを間引く 2 つの方法を合わせた早送り音声再生を図 11 を用い説明する。

【0006】 図 11 において 101 は A/D 変換器であり、102 は前記 A/D 変換器 101 で変換した音声データを、ある時間単位毎に分割する音声フレーム生成手段

である。以後このようにある時間毎に分割された音声信号を音声フレームと呼ぶ。

【0007】 103 は前記音声フレーム生成手段 102 で生成した音声フレームを入力とし、音声フレームの振幅レベルの絶対値を検知する振幅レベル検知手段である。104 は前記振幅レベル検知手段 103 で検知した振幅レベルの絶対値を入力とし、あらかじめ設定したしきい値レベルとの比較を行い、音声フレームの選択を行うしきい値比較手段である。105 は前記しきい値比較手段 104 によって選択された音声フレームを時系列に蓄積する一時バッファ手段である。106 は前記一時バッファ手段 105 に蓄積された音声フレームを管理する一時バッファ管理手段である。107 は前記一時バッファ手段 105 からの音声信号を入力とする D/A 変換器である。

【0008】 次に具体的に 1.25 倍速度の早送り音声信号の生成方法を例にとって説明する。

【0009】 まず入力音声信号を A/D 変換器 101 によってデジタル変換し、音声フレーム生成手段 102 によって、音声ピッチ周期に近い 20 ミリ秒毎に分割し、音声フレームを生成する。

【0010】 生成された音声フレームを次の振幅レベル検知手段 103 に入力し、音声フレームの振幅レベルの絶対値を検知する。

【0011】 次にしきい値比較手段 104 へ音声フレームの振幅レベルの絶対値を入力し、しきい値レベルと振幅レベルの絶対値との比較を行い、音声フレームが有音声の音声フレームであるか無音声の音声フレームであるかの判定を行う。しきい値レベルは無音状態に近い値設定する。判定は、しきい値レベルよりも振幅レベルの絶対値が大きい場合、有音声の音声フレームと判定し、小さい場合には無音声の音声フレームと判定する。

【0012】 このようにして、有音声の音声フレームと判定された音声フレームがしきい値比較手段 104 によって選択され、時系列に次の一時バッファ手段 105 へ蓄積されていく。蓄積される音声フレーム数は時間の経過とともに増加していく。

【0013】 一時バッファ管理手段 106 では、音声フレーム生成手段 102 で生成した音声フレームの数を受け、一時バッファ手段 105 に蓄積された音声フレーム数を調整する。

【0014】 ここでバッファ管理手段 106 による音声フレーム数の調整方法を説明する。説明のため一時バッファ手段 105 の容量は音声フレーム生成手段 102 で生成される音声フレームの大きさの 10 倍、つまり音声フレーム 10 フレーム分の容量を持つものとする。

【0015】 1.25 倍速度の早送り音声再生を実現するためには、時間軸方向に 8/10 の圧縮を行う必要がある。このため一時バッファ管理手段 106 によって一時バッファ手段 105 に蓄積した音声フレーム数を、音

声フレーム10フレーム分に相当する時間毎に8フレームに調整する必要がある。

【0016】調整時、一時バッファ手段105に音声フレームが10フレーム存在する場合は、蓄積された音声フレームの10フレーム中から、蓄積の順番の古い音声フレームから遡って2フレームを破棄する。9フレーム存在する場合には、同様に最後の1フレームを破棄する。このようにして音声フレームの内容に関係なく単純に音声フレームを間引いていく。

【0017】8フレーム存在する場合は破棄は行わず、7フレーム以下である場合は中身が無音である音声データの音声フレームを8フレームに満たす分だけ最後の音声フレームへ付加する。このようにして調整毎で常に8フレームになるように管理する。

【0018】このようにして一時バッファ手段105の音声フレーム数を調整毎に8フレームに管理することで、時間軸方向に8/10の圧縮が達成される。

【0019】そしてこの8つの音声フレームをD/A変換器107へ順に入力し、アナログ信号へ変換し再生することで、入力音声信号に対し1.25倍速度の早送り音声再生を達成できる。

【0020】ところで実際の音声信号、例えばテレビ番組などの音声信号を入力とした場合、無音であるシーンは少なく、しきい値比較手段104で無音の音声フレームと判定される音声フレームの数は少ない。

【0021】このため音声フレーム生成手段102によって生成された音声フレームは、しきい値比較手段104によって有音声の音声フレームと判定されることが多く、ほとんどが一時バッファ手段105へ蓄積される。これによりしきい値比較手段104での無音の音声フレームの破棄は期待できず、一時バッファ管理手段106による音声フレームの破棄が頻繁に起こる。そしてこの破棄された音声フレームの中に重要なキーワードが存在した場合は、聴取者にとって内容理解が困難な早送り音声再生になる。

【0022】

【発明が解決しようとする課題】テレビ音声などを入力音声信号とした場合、従来例では無音の音声フレームの存在が少ない場合に、一時バッファ管理手段106によって頻繁に音声フレームの破棄が生じる。このため破棄した音声フレーム中に重要なキーワードが存在した場合、内容理解が困難な早送り音声再生となっていた。

【0023】本発明ではかかる点に鑑み、無音の音声フレームが存在しない場合に生じる重要なキーワードを含む音声フレームの破棄による内容理解困難な早送り音声再生を改善するため、柔軟な無音、有音の音声フレームの判定手段を備えた音声信号処理装置を提供するものである。

【0024】

【課題を解決するための手段】前記目的を達成するため

に、本発明の音声信号処理装置は、音声信号を入力とする音声信号入力手段と、前記音声信号入力手段からの音声信号を入力とし、音声信号の分割を行なう音声フレーム生成手段と、前記音声フレーム生成手段によって分割した音声信号を入力とし、任意の周波数帯域幅に分割するフィルターバンクと、前記フィルターバンクによって周波数帯域分割した音声信号を入力とし、周波数帯域分割された音声信号から任意の周波数帯域を含む音声信号を選択する有音声判定対象帯域選択手段と、前記有音声判定対象帯域選択手段によって選択された周波数帯域制限の音声信号を入力とし、有音であるか無音であるかの判定を行ない、前記音声フレーム生成手段からの音声信号の選択を行なう有音声フレーム判定手段と、前記有音声判定手段によって選択された音声信号を入力とし、時系列に一時蓄積する一時バッファ手段と、前記一時バッファ手段に蓄積された音声信号を管理する一時バッファ管理手段と、前記一時バッファ手段からの音声信号を出力する音声信号出力手段を備えたことを特徴としている。

【0025】また、前記目的を達成するために、本発明の音声信号処理装置は、前記音声フレーム手段からの有音声信号を選択する有音声フレーム選択手段と、前記一時バッファ管理手段の代わりに、前記一時バッファ手段に蓄積された分割音声信号を、前記一時バッファ手段に蓄積された分割音声信号の数をもとに、分割音声信号単位に管理する一時バッファ管理手段を備えたことを特徴としている。

【0026】

【発明の実施の形態】

(実施例1)本発明の実施例1を図1、図2、図3、図4、図5を用いて説明する。

【0027】図1は本発明の音声信号処理装置の実施例1を示すブロック図である。図1において1は音声信号入力手段であり、A/D変換器を備える。2は前記音声信号入力手段1からの入力音声信号をある時間毎に分割し音声フレームを生成する音声フレーム生成手段であり、3は前記音声フレーム生成手段2からの音声フレームを入力とし、複数の周波数帯域へ分割するフィルターバンクであり、4は前記フィルターバンク3からの周波数帯域分割された音声フレームを入力とし、有音声判定の対象となる周波数帯域を含む周波数帯域制限の音声信号を選択する有音声判定対象帯域選択手段であり、5は前記有音声判定対象帯域選択手段4で選択した周波数帯域制限の音声信号を入力とし、有音、無音の判定を行なう有音声フレーム判定手段である。6は前記音声フレーム生成手段2で生成した音声フレームを時系列に蓄積する一時バッファ手段であり、7は前記一時バッファ手段7に蓄積された音声フレームを、前記有音声フレーム判定手段での判定結果と前記一時バッファ手段に蓄積された音声フレームの数をもとに、音声フレーム単位に管理する

一時バッファ管理手段であり、8は前記一時バッファ手段6からの音声信号を入力とする音声信号出力手段であり、D/A変換器を備える。

【0028】次に例として1.25倍速度の早送り音声再生を達成するまでの動作を説明する。

【0029】入力音声信号は音声信号入力手段1のA/D変換器によってデジタル信号へ変換され、次の音声フレーム生成手段2へ入力される。音声フレーム生成手段2では、図2に示すように入力された音声信号を音声ピッチ周期に近い20ミリ秒毎に分割し、音声フレームを生成する。

【0030】次に音声フレーム生成手段2で生成した音声フレームをフィルタバンク3へ入力する。フィルタバンク3では、図3に示すような帯域分割フィルタを複数備えたフィルタバンクを用い入力した音声フレームを複数の周波数帯域へ分割する。

【0031】フィルタバンク3で帯域分割された音声信号を入力とする有音声判定対象帯域選択手段4では、図3に示すように、人の声の再現性を重視し100Hzから200Hzの周波数を含む周波数帯域幅の音声信号を選択し、有音声フレーム判定手段5へ入力する。

【0032】有音声フレーム判定手段5では、図4に示すように、帯域制限された音声信号のスペクトル振幅レベルとしきい値レベルとの比較を行う。しきい値の設定は、スペクトル振幅レベルが0に近いレベルに設定する。

【0033】比較を行い、しきい値レベル以上のスペクトル振幅レベルのサンプルの個数を判定スコアとして算出する。この判定スコア値の大きさを参照することによって無音、有音の状態を見きわめることができる。図4に示す例では判定スコアは12になる。

【0034】以上のようにして有音声フレーム判定手段5で判定スコアが算出され、一時バッファ手段6に順次音声フレーム生成手段2からの音声フレーム出力が蓄積されていく。

【0035】次に一時バッファ手段6に蓄積した音声フレームを、一時バッファ管理手段7によって、ある時間毎にフレーム数を調整する方法について説明する。

【0036】一時バッファ管理手段7は、音声フレーム生成手段2で生成された音声フレーム数を受けて調整のタイミングを決定する。

【0037】本実施例1では一時バッファ手段6の容量である音声フレーム、10フレーム分の時間毎で調整を行う場合を例として説明する。

【0038】図5は10フレームの音声フレームが一時バッファ手段6に蓄積されている様子を示している。

【0039】各音声フレームには、有音声フレーム判定手段5で算出された判定スコアと、一時バッファ管理手段7によって管理されている固有のインデックスがあり、これらをまとめて一組とし、一時バッファ管理情報

として一時バッファ管理手段7によって管理されている。

【0040】図5の例では、音声フレームAからJに対し、各々小文字のaからjがインデックスとして対応している。各音声フレームに対応する判定スコア値は図5に示すとおりである。

【0041】一時バッファ管理手段7は、一時バッファ管理情報内の各音声フレームに対する判定スコア値を参照し、判定スコア値の大きい順に、図5のように一時バッファ管理情報の並び替えを行う。ただし、判定スコア値が同値の場合は先に一時バッファ手段6に蓄積された音声フレームを優先し並び替えを行う。

【0042】1.25倍速度の早送り音声再生を達成するためには、一時バッファ手段6内の音声フレーム数を10フレームから8フレームへ調整しなくてはならない。このため、一時バッファ管理情報の並び替えの後、図5に示すように判定スコア値の大きいものから順に、8つの音声フレームを選択し、これらを一時バッファ手段6に残し、残りの2フレームを破棄する。図5の例ではインデックスがdとgである2つの音声フレームD、Gが破棄されている。破棄した後に図5に示すように、一時バッファ手段6に残った音声フレームを次の音声信号出力手段8へ入力していく。

【0043】このように調整された一時バッファ手段6内の音声フレームを、順に音声出力手段8のD/A変換器へ入力し出力音声信号を得て、これを再生することにより1.25倍速度の早送り音声再生が達成できる。

【0044】従来例では一時バッファ手段6に蓄積された音声フレームの破棄は、音声フレームの内容に関係なく、蓄積順の古い、つまり時間的に新しい音声フレームから順に破棄していた。このため破棄した音声フレームに重要なキーワードが存在した場合、内容理解が困難な早送り音声再生になっていた。

【0045】本発明では、一時バッファ手段6に蓄積された音声フレームの判定スコア値の最も小さい、つまり無音の音声フレームにより近いものから順に音声フレームを破棄するため、従来例と比べると重要なキーワードを含む音声フレームの破棄が少なくなる。このため内容理解しやすい早送り音声再生が可能となる。

【0046】尚、本実施例1では有音声判定対象帯域手段4において、100Hzから200Hzの帯域を含む音声信号の選択を行ったが、着目する音声信号に応じて複数の帯域分割音声信号の選択を行い、前記有音声フレーム判定手段5により判定スコア値を算出することによって、同様な早送り音声再生が可能である。

【0047】また本実施例1では音声フレームの長さを20ミリ秒にしたが、既知な技術であるケプストラム法などを用いて、音声ピッチの周期を計算し、この長さで音声フレームを生成することでも同様な早送り音声再生が可能である。

【0048】また本実施例1では1.25倍速度の早送り音声再生について説明したが、一時バッファ管理手段7での音声フレームの調整の周期、破棄する音声フレームの数を調整することで、同様に任意の速度の早送り音声再生が可能である。

【0049】また本実施例1では有音声フレーム判定手段での判定スコア値の算出を、しきい値以上のスペクトル振幅を持つサンプルの個数としたが、逆にしきい値以下のものの個数を判定スコアとしても同様な早送り音声再生が可能である。

【0050】(実施例2)本発明の実施例2を図6、図7、図8、図9、図10を用いて説明する。

【0051】図6において図1と同様の機能を有するものは同一の番号を付けて説明を省略する。

【0052】図6における9は、有音声フレーム判定手段5によって算出された判定スコア値を参照し、音声フレーム生成手段2からの音声フレームデータの選択を行い、無音の音声フレームの場合は破棄を行い、有音の音声フレームの場合は一時バッファ手段6へ音声フレームを選択する有音声フレーム選択手段である。

【0053】図6における10は、有音声フレーム判定手段5からの判定スコア値、有音声フレーム選択手段9で破棄した音声フレーム数、一時バッファ手段6から蓄積した音声フレーム数を受け、一時バッファ手段に蓄積された音声フレームデータを音声フレーム単位で管理する一時バッファ管理手段である。

【0054】実施例1では、音声フレームの破棄を実施例1で説明した一時バッファ管理手段のみによって行っていた。本実施例2では、有音声フレーム選択手段9で無音と判定される音声フレームをあらかじめ破棄しておき、さらに図6の一時バッファ管理手段10によって一時バッファ手段6中の音声フレーム数を調整することで早送り音声再生を達成する。実施例1と同様に1.25倍速度の早送り音声再生を例にとって説明を行う。

【0055】音声信号入力手段1へ入力音声信号を入力し、有音声フレーム判定手段5によって判定スコア値を算出するまでは、前記の実施例1で説明したものと同様であるので説明を省略する。

【0056】有音声フレーム選択手段9は、有音声フレーム判定手段5によって算出された判定スコア値を参照し、図7に示すような音声フレームの選択を行う。判定スコア値が正值の場合、つまり有音の音声フレームとして判定できる場合は、音声フレームを次の一時バッファ手段6へ入力し、判定スコア値が0の場合、つまり無音の音声フレームとして判定できる場合は破棄を行う。このようにして音声フレームの選択を行う。

【0057】また、破棄した音声フレーム数を一時バッファ管理手段10へ入力する。このようにして、選択された音声フレームは一時バッファ手段6へ時系列に蓄積されていく。

【0058】次に一時バッファ管理手段10による一時バッファ6中の音声フレームの調整方法について図8、図9、図10を用いて説明する。

【0059】調整は、あらかじめ設定した音声フレーム数が、一時バッファ手段6に蓄積された時点で始まる。

【0060】ここでは例として、音声フレームが20フレーム蓄積された時点で調整を行う場合をとりあげる。尚、調整開始時に一時バッファ手段6の音声フレーム数、20フレームに対して、有音声フレーム選択手段9による破棄した音声フレーム数が5フレームであれば、4対1の割合となり時間軸方向に8/10の圧縮ができ、この20フレームを次の音声信号出力手段に入力し再生すれば1.25倍速度の早送り音声再生が達成される。つまり、一時バッファ手段に残す音声フレーム数と、最終的に破棄される音声フレーム数の比を4対1に調整すれば、1.25倍速度の早送り音声再生が達成されることになる。

【0061】一時バッファ手段6に音声フレームが20フレーム蓄積されるまでに、有音声フレーム選択手段によって破棄された音声フレームの数は、音声フレームの内容によって一定ではない。調整時での一時バッファ手段中の音声フレーム数と、有音声フレーム選択手段によって破棄された音声フレーム数の状態は、図8、図9に示す2通りが考えられる。

【0062】図8は、一時バッファ手段に20フレームの音声フレームが蓄積されており、有音声フレーム選択手段によって7フレームの音声フレームが無音の音声フレームとして破棄された場合を示している。

【0063】図9は、同様に一時バッファ手段に20フレームの音声フレームが蓄積されており、音声フレーム選択手段によって3フレームの音声フレームが無音の音声フレームとして破棄された場合を示している。

【0064】図8、図9の最小無音声フレーム数とは、一時バッファ手段中の20フレームに対する4分の1の値である定数値の5フレームを示している。

【0065】まず図8に示す場合の一時バッファ管理手段による調整の流れを図10を用いて説明する。

【0066】図10は一時バッファ管理手段による調整の流れを示したもので、20フレーム分の音声フレームが一時バッファ手段に蓄積された時点で調整が開始される。図8に示す場合ではステップ201、ステップ202、ステップ203、ステップ204が実行される。

【0067】まずステップ201では、前回の調整時にステップ204でセットした無音声フレーム数の余りを現在の無音声フレーム数へ加算する。

【0068】ここで無音声フレーム数の余りとは、図8の場合であれば、最小無音声フレーム数を超えた分の音声フレーム数を示しており、図8に示している現在の無音声フレーム数7フレームは、前回の無音声フレーム数の余りを含めたフレーム数である。また図8では次回に

加算される無音声フレーム数の余りは2フレームとなる。尚、無音声フレーム数の初期値は0フレームである。

【0069】ステップ201によって、前回の無音声フレーム数の余りを含めた無音声フレーム数を決定し、次のステップ202では無音声フレーム数と、最小無音声フレーム数との比較を行う。図8の場合は、比較は真値となり、次のステップ203へ進む。

【0070】ステップ203では、一時バッファ手段に蓄積された20フレームの音声フレーム全てを音声信号出力手段へ入力する。一時バッファ手段中の音声フレームの破棄は一切行わない。

【0071】次のステップ204では、次回の調整時のステップ201で加算する無音声フレーム数の余りを算出する。図8においては2フレームとなる。

【0072】図8に示すように、有音声フレーム選択手段によって破棄された音声フレーム数が最小無音声フレーム数以上の場合、最小無音声フレーム5フレームに対し、一時バッファ手段中の20フレームの音声フレームをそのまま音声信号出力手段へ入力し再生することで1.25倍速度の早送り音声再生が達成される。この場合の音声は、判定スコア値が0の音声フレームだけの廃棄を行った1.25倍速度音声となる。

【0073】次に図9の有音声フレーム選択手段によって破棄された音声フレーム数が最小無音声フレーム数より少ない場合を説明する。尚、ステップ201、ステップ202までは同様であるため説明を省略し、ステップ205からの流れを説明する。

【0074】図9は最小無音声フレーム数が、無音声フレーム数より多い場合であり、無音声フレーム数は3フレームとなっている。このまま一時バッファ手段の20フレームの音声フレームを音声信号出力手段へ入力しても、現在の無音声フレーム数3フレームに対し20フレームの音声フレームの再生となるため、1.25倍速度に満たない早送り音声再生となる。このため、ステップ205では最小無音声フレーム数に満たない数の音声フレームを新たに処理し、有音声フレーム選択手段によって、破棄または一時バッファ手段へ音声フレームを蓄積する。図9の場合は、2フレーム中、1フレームが破棄され無音声フレーム数が3から4になり、1フレームが一時バッファ手段へ蓄積され21フレームの音声フレームが一時バッファ手段に蓄積された場合を示している。

【0075】ステップ205で新規に音声フレームを処理した後に、ステップ206では新たに一時バッファ手段に蓄積された音声フレーム数分を、一時バッファ手段に蓄積された音声フレームから破棄する。図9の場合は、新たに蓄積した音声フレーム数は1フレームであるので、1フレーム分の音声フレームを破棄する。そしてこの破棄の方法は前期実施例1で図5を用いて説明した方法を用い、判定スコア値をもとに1フレーム分を破棄

する。

【0076】ステップ206で音声フレームの破棄を行った後は20フレームの音声フレームが一時バッファ手段に残されることになり、この20フレームの音声フレームを音声信号出力手段へ入力する。

【0077】最後にステップ207で次回に加算する無音声フレームの余りを0にセットする。

【0078】図9の場合、有音声フレーム選択手段にて破棄される音声フレームが少ない場合、実施例1で図5を用いて説明した判定スコア値をもとにした音声フレームの破棄を行い、図8の場合と同様に、結果的に破棄した総音声フレーム数5フレームに対し、20フレームの音声フレームを音声信号出力手段へ入力し再生することで、1.25倍速度の早送り音声再生が達成できる。

【0079】前記の実施例1の例では、判定スコア値をもとに、あらかじめ決められた数の音声フレームの破棄を行い可変速度音声再生を達成していた。これは単純に時間的に新しい音声フレームを破棄する従来の方法よりも、音声フレームの内容、つまり判定スコア値をもとに破棄する音声フレームを決定する点で従来の例と比較するとキーワードの欠落の少ない早送り音声再生が達成できる。

【0080】尚、本実施例2では1.25倍速度の早送り音声再生について説明したが、実施例1と同様に一時バッファ管理手段7での調整する音声フレーム数を変えることで任意の速度の早送り音声再生が可能である。

【0081】また、図10で、ステップ203とステップ204を入れ替えてもどのような早送り音声再生が達成でき、またステップ207とステップ208を入れ替えてもどのような早送り音声再生が達成できる。

【0082】

【発明の効果】以上説明したように、柔軟な音声フレームの破棄を行う一時バッファ管理手段を備えることによって、重要なキーワードを含んだ音声信号の欠落を極力抑え、聞き取りやすい早送り音声再生を達成できる。

【図面の簡単な説明】

【図1】実施例1における音声処理装置の構成を示すブロック図

【図2】図1における音声フレームの生成過程の示す模式図

【図3】図1におけるフィルターバンク3、有音声対象帯域選択手段4での動作を示す模式図

【図4】図1における判定スコア値の説明図

【図5】図1における一時バッファ管理手段7での音声フレーム数の調整を示す模式図

【図6】実施例2における音声処理装置のブロック図

【図7】図6における有音声フレーム選択手段9での音声フレームの選択を示す模式図

【図8】図6における一時バッファ管理手段10での音声フレーム数の調整を示す模式図

【図9】図6における一時バッファ管理手段10での音声フレーム数の調整を示す模式図

【図10】図6における一時バッファ管理手段10での音声フレーム数の調整を示すフローチャート

【図11】従来例を示すブロック図

【符号の説明】

1 音声信号入力手段

\*

\* 2 音声フレーム生成手段

3 フィルターバンク

4 有音声判定対象帯域選択手段

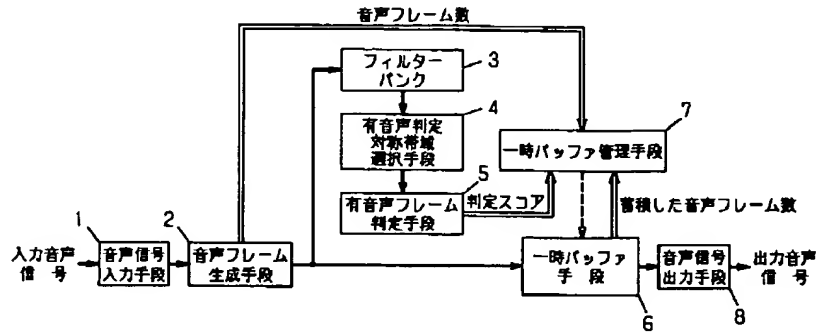
5 有音声フレーム判定手段

6 一時バッファ手段

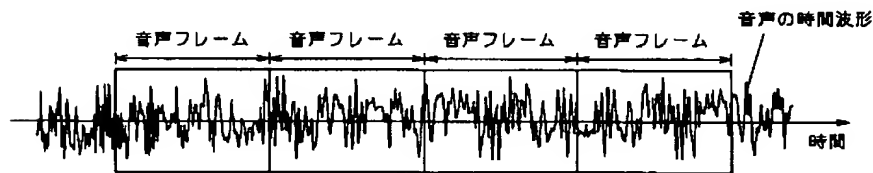
7 一時バッファ管理手段

\* 8 音声信号出力手段

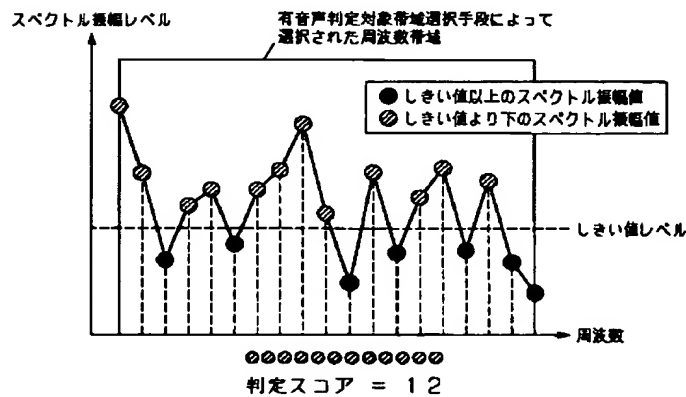
【図1】



【図2】

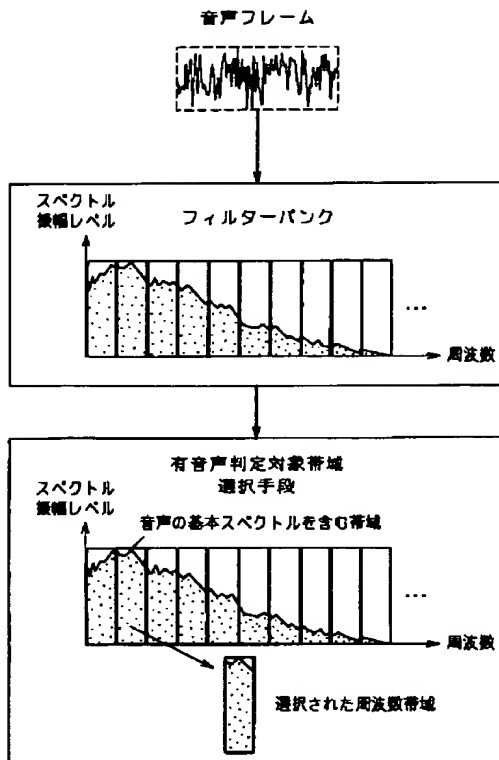


【図4】

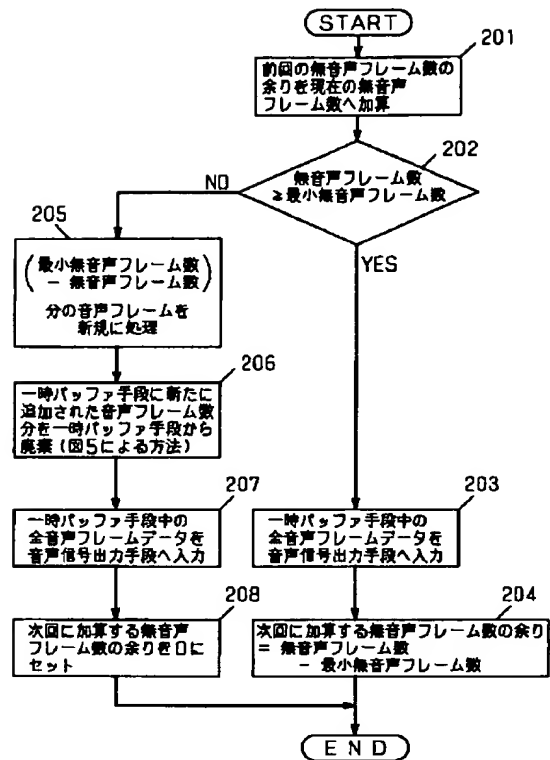




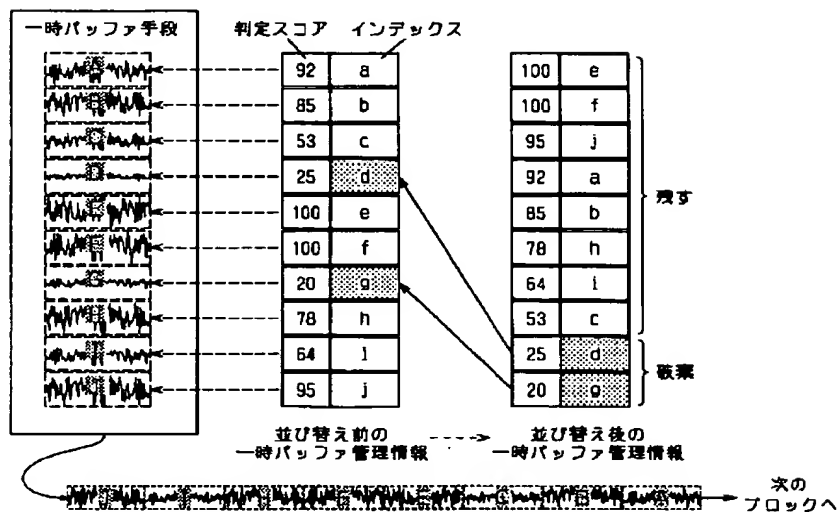
【図3】



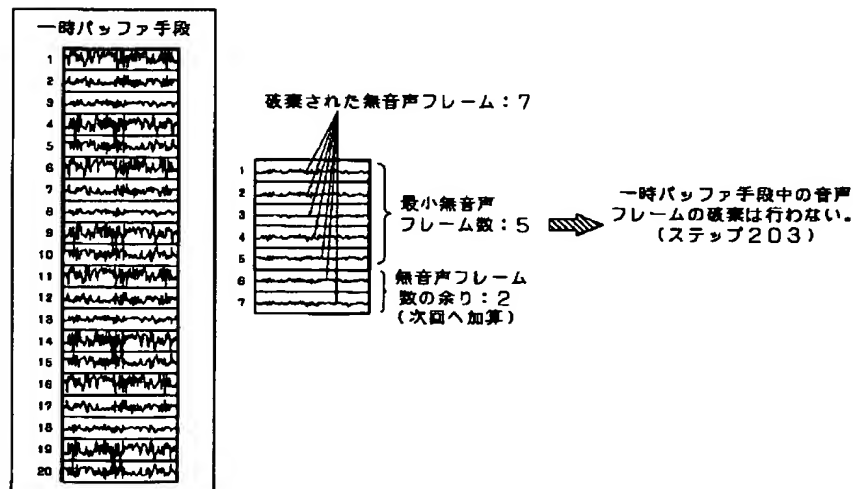
【図10】



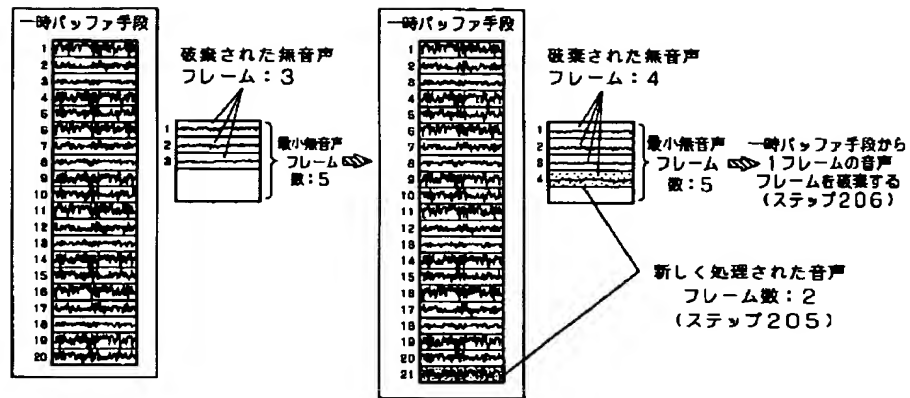
【図5】



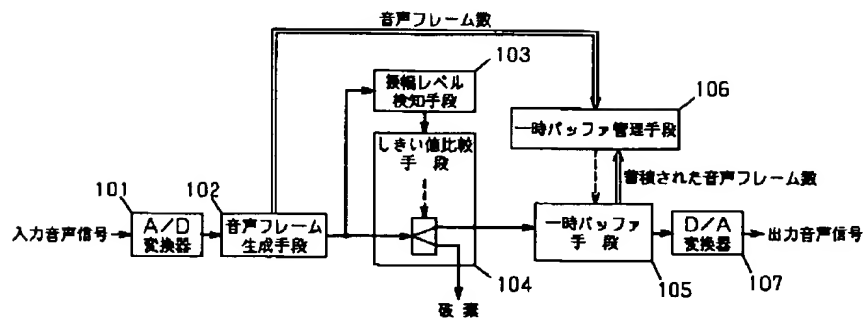
音声フレーム数



【図9】



【図11】



フロントページの続き

(72)発明者 石津 厚  
大阪府門真市大字門真1006番地 松下電器  
産業株式会社内